NASA-CR-192731

NAGW-1333

CIRSSE DOCUMENT #1

TECHNICAL REPORTS

N93-71635

IN-63-CR

153766

P-8

(NASA-CR-192731) ANALYTIC FORMULATION OF INTELLIGENT MACHINES AS NEURAL NETS (Rensselaer Polytechnic Inst.) 8 p

Unclas

Z9/63 0153766

Center for Intelligent Robotic Systems for Space Exploration

Rensselaer Polytechnic Institute
Troy, New York 12180-3590

ANALYTIC FORMULATION OF INTELLIGENT MACHINES AS NEURAL NETS

By:

G.N. Saridis
M.C. Moed

Department of Electrical, Computer and Systems Engineering
Department of Mechanical Engineering, Aeronautical Engineering & Mechanics
Rensselaer Polytechnic Institute
Troy, New York 12180-3590

CIRSSE Document #1

ANALYTIC FORMULATION OF INTELLIGENT MACHINES AS NEURAL NETS

George N. Saridis
and Michael C. Moed

Robotics and Automation Laboratories
Electrical, Computer and Systems Engineering Department
Rensselaer Polytechnic Institute
Troy, New York 12180-3590

ABSTRACT

Recent technological developments have fostered a need for the development and utilization of machines which contain enough intelligence to perform autonomous tasks in uncertain environments. Concepts drawn from the fields of Artificial Intelligence, Operations Research and Control Theory have been combined to form a unified theory which analytically describes the design and structure consisting of Organization, Coordination and Execution levels forms the architecture of the Machine under the Principle of Increasing Precision with Decreasing Intelligence from hierarchically intelligent control. This system has been formulated as a probabilistic model, where uncertainty and imprecision can be expressed in terms of entropies. The optimal strategy for decision planning and task execution by the Intelligent Machine can be found by minimizing the total entropy in the system.

This paper focuses on the design of the Organization Level of the Intelligent Machine as a Boltzmann machine, as described in current neural network literature. Since this level is responsible for planning the actions of the Machine, the problem at this tier is formulated as the construction of the right sequence of tasks or events which minimizes the entropy for the desired action. A search algorithm is presented which examines the weights and connections of the events in order to efficiently find the desired action sequence.

1. INTRODUCTION

In our present technological society, there is a major need to build machines that would execute intelligent tasks operating in uncertain environments with minimum interaction with a human operator. Although some designers have built smart robots, utilizing heuristic ideas, there is no systematic approach to design such machines in an engineering manner.

Recently, cross-disciplinary research from the fields of computers, systems, AI and information theory has served to set the foundations of the emerging area of the Design of Intelligent Machines (Saridis, Stephanou 1977).

Since 1977 Saridis has been developing a novel approach, defined as Hierarchical Intelligent Control, designed to organize, coordinate and execute anthropomorphic tasks by a machine with minimum interaction with a human operator. This approach utilizes analytical (probabilistic) models to describe and control the various functions of the Intelligent Machine structured by the intuitively defined principle of Increasing Precision with Decreasing Intelligence (IPDI) (Saridis 1979).

This principle, even though resembles the managerial structure of organizational systems (Levis 1988), has been derived on an analytic basis by Saridis (1988). The impact of this work is in the engineering design of intelligent robots, since it provides analytic techniques for universal production (blueprints) of such machines.

The purpose of the paper is to derive analytically a Boltzmann machine suitable for optimal connection of nodes in a neural net (Fahlman, Hinton, Sejnowski, 1985). Then this machine will serve to search for the optimal design of the Organization level of an Intelligent Machine.

In order to accomplish this, some mathematical theory of the intelligent machines will be first outlined. Then some definitions of the variables associated with the principle, like machine intelligence, machine knowledge, and precision will be made. A list of such definitions is given in the section that follows. (Saridis, Valavanis 1988). Then a procedure to establish the Boltzmann machine on an analytic basis will be presented and illustrated by an example in designing the organization level of an Intelligent Machine.

2. THE MATHEMATICAL THEORY OF INTELLIGENT CONTROLS

In order to design intelligent machines that require for their operation control system with intelligent functions such as simultaneous utilization of a memory, learning, or multilevel decision making in response to "fuzzy" or qualitative commands, Intelligent Controls have been developed by Saridis (1977, 1983). They utilize the results of cognitive systems research effectively with various mathematical programming control techniques (Birk & Kelley, 1981).

The theory of Intelligent Control systems, proposed by Saridis (1979) combines the powerful high-level decision making of the digital computer with advanced mathematical modeling and synthesis techniques of system theory with linguistic methods of dealing with imprecise or incomplete information. This produces a unified approach suitable for the engineering needs of the future. The theory may be thought of as the result of the intersection of the three major disciplines of Artificial Intelligence, Operations Research and Control Theory. This research is aimed to establish Intelligent Controls as an engineering discipline, and it plays a central role in the design of Intelligent Autonomous Systems.

Intelligent control can be considered as a fusion between the mathematical and linguistic methods and algorithms applied to systems and processes. In order to solve the modern technological problems that require control systems with intelligent functions such as simultaneous utilization of a memory, learning, or multilevel decision making in response to "fuzzy" or qualitative commands. Intelligent Control is the process of implementation of an Intelligent Machine and would require a combination of "machine intelligent functions" for task organization purposes with system theoretic methods for their execution.

The control intelligence is hierarchically distributed according to the Principle of Increasing Precision with Decreasing Intelligence (IPDI), evident in all hierarchical management systems. They are composed of three basic levels of controls even though each level may contain more than one layer of tree-structured functions (Figure 1):

1. The organization level.
2. The coordination level.
3. The execution level.

The Organization Level is intended to perform such operations as planning and high level decision making from long term memories. It may require high level information processing such as the knowledge based systems encountered in Artificial Intelligence. These require large quantities of knowledge processing but require little or no precision.

The functions involved in the upper levels of an intelligent machine are imitating functions of human behavior and may be treated as elements of knowledge-based systems. Actually, the activities of planning, decision making, learning, data storage and retrieval, task coordination, etc. may be thought of as knowledge handling and management. Therefore, the flow of knowledge in an intelligent machine may be considered as the key variable of such a system.

Knowledge flow in an intelligent machine's organization level represents respectively:

1. Data Handling and Management.
2. Planning and Decision performed by the central processing units.
3. Sensing and Data Acquisition obtained through peripheral devices.
4. Formal Languages which define the software.

Subjective probabilistic models or fuzzy sets are assigned to the individual functions. Thus, their entropies may be evaluated for every task executed. This provides an analytical measure of the total activity.

Artificial Intelligence methods also applicable for the processing of knowledge and knowledge rates of the organization level of an intelligent machine have been developed by Meystel (1985) and his colleagues.

The Coordination Level is an intermediate structure serving as an interface between the organization and execution level.

It is involved with coordination, decision making and learning on a short term memory, e.g., a buffer. It may utilize linguistic decision schemata with learning capabilities defined in Saridis and Graham (1984), and assign subjective probabilities for each action. The respective entropies may be obtained directly from these subjective probabilities.

The Execution Level executes the appropriate control functions. Its performance measure can also be expressed as an entropy, thus unifying the functions of an "intelligent machine".

Optimal control theory utilizes a non-negative functional of the states of a system in the states space, and a specific control from the set of all admissible controls to define the performance measure for some initial conditions $(x(t), t)$, representing a generalized energy function. Minimization of the energy functional yields the desired control law for the system.

For an appropriate density function $p(x, u(x,t), t)$ satisfying Jaynes' Maximum entropy principle (1957), it was shown by Saridis (1988) that the entropy for a particular control action $u(x,t)$.

$$H(u) = \int_{\Omega_x} p(x, u(x,t), t) \ln p(x, u(t), t) dx$$

is equivalent to the expected energy or cost functional of the system. Therefore, minimization of the entropy $H(u)$ yields the optimal control law of the systems.

This statement establishes equivalent measures between information theoretic and optimal control problems and unifies both information and feedback control theories with a common measure of performance. Entropy satisfies the additive property, and any system composed of a combination of such subsystems can be optimized by minimizing its total entropy. Information theoretic methods based on entropy may apply (Conant 1976).

Since all levels of a hierarchical intelligent control can be measured by entropies and their rates, then the optimal operation of an "intelligent machine" can be obtained through the solution of mathematical programming problems.

The various aspects of the theory of hierarchically intelligent controls may be summarized as follows:

The theory of intelligent machines may be postulated as the mathematical problem of finding the right sequence of decisions and controls for a system structured according to the principle of increasing precision with decreasing intelligence (constraint) such that it minimizes its total entropy.

The above analytic formulation of the "intelligent machine problem" as a hierarchically intelligent control problem is based on the use of entropy as a measure of performance at all the levels of the hierarchy. It has many advantages because of the tree-like structure of the decision making process, and brings together functions that belong to a variety of disciplines.

3. KNOWLEDGE FLOW AND THE PRINCIPLE OF IPDI

The concept of entropy used in this paper may be generalized if one introduces theory of evidence for the cases that Intelligent Machines are endowed with judgment, a very human property.

The general concepts of Intelligent Control Systems are the fundamental notions of Machine Intelligence, Machine Knowledge, its Rate and Precision. The following definitions are useful in order to derive the principle of IPDI.

Def. 1 **Machine Knowledge is defined to be the structured information acquired and applied to remove ignorance or uncertainty about a specific task pertaining to the Intelligent Machine.**

Knowledge is a cumulative quantity accrued by the machine and cannot be used as a variable to execute a task. Instead, the Rate of Machine Knowledge is a suitable variable.

Def. 2 **Rate of Machine Knowledge is the flow of knowledge through an Intelligent Machine.**

Intelligence is defined by the American Heritage Dictionary of the English Language (1969) as: Intelligence is the capacity to acquire and apply knowledge.

In terms of Machine Intelligence, this definition may be modified to yield:

Def. 3 **Machine Intelligence (MI) is the variable (source) which operates on a data-base (DB) of events to produce flow of knowledge (RK)**

One may directly apply the Law of Partition of Information Rates of Conant (1976) to analyze the functions of intelligence within the activities of an Intelligent Control System.

On the other hand, one may define Precision as follows:

Def. 4 **Imprecision is the uncertainty of execution of the various tasks of the Intelligent Machine.**

and

Def. 5 **Precision is the complement of Imprecision, and represents the complexity of a process.**

Analytically, the above relations may be summarized as follows:

Knowledge (K) representing a type of information may be represented as

$$K = -\alpha - \ln p(K) = (\text{Energy}) \tag{1}$$

where $p(K)$ is the probability density of Knowledge.

From equation (1) the probability density function $p(K)$ satisfies the following expression in agreement with Jaynes' principle of Maximum Entropy (1957):

$$p(K) = e^{-\alpha - K}; \quad \alpha = \ln \int_X e^{-K} dx \tag{2}$$

The Rate of Knowledge R which is the main variable of an intelligent machine with discrete states is

$$R = \frac{K}{T} = (\text{Power})$$

It was intuitively thought (Saridis 1983), that the Rate of Knowledge must satisfy the following relation which may be thought of expressing the principle of Increasing Precision with Decreasing Intelligence

$$(MI):(DB) \longrightarrow (R) \tag{3}$$

A special case with obvious interpretation is, when R is fixed, machine intelligence is largest for a smaller data base e.g. complexity of the process. This is in agreement with Vamos' theory of Metalanguages (1986).

It is interesting to notice the resemblance of this entropy formulation of the Intelligent Control Problem with the ε-entropy formulation of the metric theory of complexity originated by Kolomogorov (1956) and applied to system theory by Zames (1979). Both methods imply that an increase in Knowledge (feedback) reduces the amount of entropy (ε-entropy) which measures the uncertainty involved with the system.

An analytic formulation of the above principle has been derived from simple probabilistic relation among the Rate of Knowledge, Machine Intelligence and the Data Base of Knowledge. The entropies of the various functions come naturally into the picture as a measure of their activities.

4. THE DESIGN OF THE ORGANIZATION LEVEL OF AN INTELLIGENT MACHINE AS A BOLTZMANN MACHINE

In the current literature of parallel architectures for Machine Intelligence, the Boltzmann machine represents a powerful, neural network based architecture that allows efficient searches to optimally obtain the combination of certain hypotheses of input data and constraints (Fahlman, Hinton, Sejnowski 1985).

The Boltzmann architecture may be interpreted as the machine that searches for the optimal interconnection of several nodes (neurons) representing different primitive events in order to produce a string defining an optimal task. Such a device may prove extremely useful for the design of the Organization Level of an Intelligent Machine. (Saridis, Valavanis 1988). (Figure 2)

We associate the state of each node with a binary random variable $x_i = \{0,1\}$, with a priori probabilities $p(x_i = 1) = p_i$, $p(x_i = 0) = 1 - p_i$, where 1 represents the firing of neuron i, and 0 indicates neuron i idle. The state vector of the network, $X = \{x_1, x_2, \ldots, x_i, \ldots, x_n\}$ is an ordered set of 0's and 1's describing the state of the machine in terms of firing/idle nodes, for an n node machine. The neurons of the machine can be visible, or hidden (Hinton, Sejnowksi 1986). It is possible to extract the string of primitive events representing the optimal task by examining the state vector of the visible nodes in the network in steady state response to a given input.

5. ENTROPY AS A MEASURE OF UNCERTAINTY

Entropy is used as a measure of uncertainty in the intelligent machine. The entropy manifests itself in the interaction and interconnection of nodes in the network. We can define the energy of flow of knowledge from node j to i by:

$$R_{ij} = \frac{1}{2} w_{ij} x_i x_j \tag{4}$$

where $w_{ii} = 0$, which is analogous to Hopfield's neurological model (Hopfield 1982).

The Probability of Knowledge flow from node j to i is:

$$p(R_{ij}) = e^{-\alpha_i - \frac{1}{2} w_{ij} x_i x_j} \quad \text{from (2)} \tag{5}$$

where $\alpha_i \geq 0$ is the probability normalizing factor.

The Entropy of Knowledge Flow from node j to i is:

$$H(R_{ij}) = -p(R_{ij}) ln\{p(R_{ij})\} \tag{6}$$

or:

$$H(R_{ij}) = (\alpha_i + \frac{1}{2} w_{ij} x_i x_j) e^{-\alpha_i - \frac{1}{2} w_{ij} x_i x_j} \tag{7}$$

If x_i or $x_j = 0$:

$$H(R_{ij}) = \alpha_i e^{-\alpha_i} \tag{8}$$

which can be called the Threshold Node Entropy of node i.

If x_i and $x_j = 1$:

$$H(R_{ij}) = (\alpha_i + \frac{1}{2} w_{ij})(e^{-\alpha_i - \frac{1}{2} w_{ij}}) \tag{9}$$

Similarly, the Entropy of the Flow of Knowledge into node i is:

$$H(R_i) = \left(\alpha_i + \frac{1}{2}\sum_j w_{ij} x_i x_j\right)\left(e^{-\alpha_i - \sum_j w_{ij} x_i x_j}\right) \tag{10}$$

$$H(R_i) = \left(\alpha_i + \frac{1}{2}\sum_j w_{ij} x_i x_j\right)\left(e^{-\alpha_i} \prod_j e^{-\frac{1}{2} w_{ij} x_i x_j}\right) \tag{11}$$

$$H(R_i) = \left(\alpha_i + \sum_j R_{ij}\right)\left(e^{-\alpha_i}\right)\left(\prod) j e^{-R_{ij}}\right) \tag{12}$$

The Entropy of Flow of Knowledge in the Intelligent Machine is:

$$H(R) = \left[\sum_i (\alpha_i + \frac{1}{2}\sum_j w_{ij} x_i x_j)\right] \prod_i \left[e^{-\alpha_i} \prod_j e^{-\frac{1}{2} w_{ij} x_i x_j}\right] \tag{13}$$

With the uncertainty of the network measured, the reduction of the Entropy of Knowledge Flow must be examined. Each node i has a Threshold Node Entropy associated with it, as shown in (8). This Threshold Node Entropy can also be found by setting $x_j = 0$ for all nodes j ($j \neq 1$) in (11). The Threshold Node Entropy is the entropy of a node when no knowledge flows into the node. It is necessary to assert some nodes $j_1, \ldots, j_k$ which connect to node i such that the Entropy of the Flow of Knowledge into node i is less than the Threshold Node Entropy of i, in order to reduce the entropy in the machine. This corresponds to the reduction of uncertainty in the machine through knowledge acquisition.

Form (8) and (9) we can show that the Entropy of Knowledge Flow from node j to node i will be less than the Threshold Node Entropy of node i under the following condition:

$$\alpha_i e^{-\alpha_i} > \left(\alpha_i + \frac{1}{2} w_{ij}\right)\left(e^{-\alpha_i} - \frac{1}{2} w_{ij}\right)$$

$$\alpha_i > \left(\alpha_i + \frac{1}{2} w_{ij}\right)\left(e^{-\frac{1}{2} w_{ij}}\right)$$

Then:

$$\alpha_i > \frac{\frac{1}{2} w_{ij}}{e^{\frac{1}{2} w_{ij}} - 1} \tag{14}$$

So if $x_i, x_j = 1$, both nodes will be asserted and the Entropy Flow from Node j to node i will be less than the Threshold Node Entropy of i when (14) holds. For example, given:

$$w_{ij} = 0.8 \ , \ \text{then} \ \frac{\frac{1}{2} w_{ij}}{e^{\frac{1}{2} w_{ij}} - 1} = 0.813$$

So if $\alpha_i > 0.813$, the local entropy will be reduced by asserting x_i, x_j.

Similarly, we can consider the total Entropy of the Flow of Knowledge into node i from all other nodes. From (11) we can derive conditions such that the assertion of nodes connected to node i will reduce the Entropy at node i from its Threshold Node Entropy:

$$H(R_i) = \left(\alpha_i + \frac{1}{2}\sum_j w_{ij} x_i x_j\right) e^{-\alpha_i} \prod_j e^{-\frac{1}{2} w_{ij} x_i x_j}$$

Assuming $x_i = 1$:

Then

$$\alpha_i > \frac{\frac{1}{2}\sum_j w_{ij} x_j}{1 - \prod_j e^{-\frac{1}{2} w_{ij} x_j}} \tag{15}$$

In similar ways, the Threshold Net Entropy can be determined, and Net Entropy reduction criteria developed from (13).

6. SEARCH TECHNIQUES FOR THE INTELLIGENT MACHINE

Two random search techniques are compared here which may be used to find the minimum entropy in the Organization Level of an intelligent machine. By examining the active visible neurons in the minimum entropy state of the network, one can determine the sequence of primitive events which produce a string defining an optimal task for an intelligent machine. The techniques presented here allow escape from local entropy minima, which lead to incorrect task decisions, by randomly selecting states while searching for the global entropy minimum.

6.1 Simulated Annealing

One random search technique commonly used to find the global minimum cost in a Boltzmann Machine is Simulated Annealing. This technique simulates the annealing process of metal by probabilistically allowing uphill steps in a state–dependent cost function while finding the global cost minimum, or ground state. The algorithm allows control of the search randomness by a user specified parameter, T. In true metal annealing, this cost function is the Energy of the system, E, and T is the annealing temperature (Kirkpatrick et al. 1983). This method can easily be adapted for finding the minimum entropy of the Organization Level of an intelligent machine.

Given is a small random change in the system state $X_i = \{x_1, x_2, \ldots, x_n\}$ to X_i' and the resulting entropy change, ΔH, if $\Delta H \leq 0$, the change is accepted. If $\Delta H > 0$, the probability the new state is accepted is:

$$p(X_{i+1} = X_i') = e^{-\Delta H / K_B T} \tag{16}$$

where K_B is the Boltzmann Constant and T is a user set parameter. By reducing T along a schedule, called the annealing schedule, the system should settle into a near–ground state as T approaches 0.

Another method for simulated annealing is discussed in (Hinton, Sejnowski 1986). Using this method, if the entropy change between X_i and X_i' is ΔH, then regardless of the previous state, accept state X_i' with probability:

$$p(X_{i+1} = X_i') = \frac{1}{1 + e^{-\Delta H / T}} \tag{17}$$

Since an intelligent machine consists of a set of binary states, it should be noted that in both of the above methods, X_i' is hamming distance 1 from X_i (Kam et al. 1985).

The process of simulated annealing escapes local minima through its probabilistic random search, and probabilistically convergences to the global cost minimum. Under certain conditions (Geman, Geman 1984). The next technique, Expanding Subinterval Random Search, probabilistically guarantees convergence within a δ neighborhood to the global minimum of a specified cost function.

6.2 Expanding Subinterval Random Search

A second technique for finding the global minimum value for a cost function for a dynamic system is Expanding Subinterval Random Search as described in (Saridis 1976). Using entropy as the cost function and given a state X_i, one may define the following random search algorithm for an appropriately selected μ.

$$X_{i+1} = \begin{cases} X_i' & \text{if} \quad H(X_i') - H(X_i) \leq 2\mu \\ X_i & \text{if} \quad H(X_i') - H(X_i) > 2\mu \end{cases} \tag{18}$$

where $H(Y)$ is the entropy induced by state $Y = (y_1, y_2, \ldots, y_n)$ and X_i' is a randomly selected state vector generated from a prespecified independent and identically distributed density function, defined by (5).

It is shown that:

$$\lim_{n \to \infty} \; Prob \; [H(X_n) - H^*_{\min} < \delta] = 1 \tag{19}$$

where $H^*_{\min}$ is the global minimum entropy of the network. The existence of $H^*_{\min}$ is proven in the cited work.

This method can be used on–line to find the global minimum entropy in the Organization Level of an intelligent machine.

7. EXPERIMENTAL RESULTS

7.1 Simulation of Search Techniques

A net was created which recognized strings of 15 bit binary numbers. The net was formulated using the standard Energy methods found in (Hinton, Sejnowski 1986). Energy was used instead of Entropy in these simulations for two reasons. First, to compare the results of this simulation to the results of simulations by other researchers, a standard measure had to be used. Second, the method for creating regions of attraction in an Entropy based net is still being investigated.

The net had three minima, corresponding to states (100101000100010, 011101010010011, 000111000000100). The respective Energy values for these three states were (0.5, 0.45, 0.2). Each simulation technique attempted to find the global Energy minimum of the net, which was 0.2. Three cases were studied which varied the depth and narrowness of the Energy well for the global minimum state.

Simulated Annealing was performed using the acceptance criteria in (17). The system was cooled in accordance with:

$$\frac{T_1(t)}{T_0} = \frac{1}{\log(10 + t)}$$

where $T_1(t)$ = temperature at time t
T_0 = initial temperature.

The net state changed in Hamming distance 1 increments.

Expanding Subinterval Random Search (Saridis 1976) was slightly modified to reinforce the probabilistic selection of node states which reduced the Energy in the net. The probability of a node being active as initially 0.5. When the Energy was reduced during search, the probability of the node being reactivated became

$$P(x_i = 1) = P(x_i = 1) + [1.0 - P(x_i = 1)] * 0.1$$

if the node was active, or

$$P(x_i = 1) = P(x_i = 1) - P(x_i = 1) * 0.1$$

if the node was inactive.

Figures 3–8 present typical simulation results. In the first experiment, the Energy well was small and wide compared to later experiments. The second experiment approximately doubles the well depth and narrowness, and the third experiment approximately triples the original values.

As one can see from the data, Simulated Annealing (SA) found the global minimum Energy in the first experiment in approximately an order of magnitude interactions faster than Expanding Subinterval Random Search (ESRS). In the second experiment, ESRS converged to the global minimum in an order of magnitude iterations quicker than SA. In the last experiment, SA did not converge to the global minimum. Also, it is shown that SA did not settle to the minimum Energy in any of the experiments, while ESRS settled in every trial.

This experiment indicates that Simulated Annealing quickly finds the global minimum if this value does not reside within a deep Energy well. Since SA searches locally for lower Energy states by changing one node value at a time, it may not find a lower value when the search must significantly climb before the minimum can be seen. Expanding Subinterval Random Search will converge to the global Energy minimum regardless of the depth and narrowness of the Energy well, but converges slower than SA when this topology does not hold.

8. CONCLUSIONS

A mathematical theory for intelligent machines was proposed and traced back to its origins. The methodology was developed to formulate the "intelligent machine", of which an intelligent robot system is a typical example, as a mathematical programming problem as using the aggregated entropy of the system as its performance measure. The levels of the machine structured according to the Principle of Increasing Precision with Decreasing Intelligence can adapt performance measures easily expressed as entropies. This work establishes an analytic formulation of the Principle, provides entropy measures for the account of the underlying activities, and integrates it with the main theory of

"Intelligent Machines". Optimal solutions of the problem of the "intelligent machine" can be obtained by minimizing the overall entropy of the system.

This formulation was proven to be applicable to the derivation and design of parallel architectures for Machine Intelligence. The Boltzmann machine was analytically derived from the definitions of knowledge flow and Jaynes' principle of maximum entropy. An analytic formulation was given to reduce the entropy due to knowledge flow between active nodes. Two techniques, Simulated Annealing and Expanding Subinterval Random Search, were described. These techniques are used to find the global minimum entropy of a Boltzmann Machine. Simulations using these search techniques were conducted using Energy as a cost function, and results indicate that ESRS converges faster than SA to a global minimum if the topology contains narrow and deep cost wells.

REFERENCES

Albus, J. S., (1975. "A New Approach to Manipulation Control: The Cerebellar Model Articulation Controller". *Transactions of ASME, J. Dynamics Systems, Measurement and Control, 97*, 220-227.

American Heritage Dictionary of the English Language, (1969).

Birk, J. R. and Kelley, R. B., (1981. "An Overview of the Basic Research Needed to Advance the State of Knowledge in Robotics". *IEEE Trans. on SMC*, SMC-11, No. 8, pp. 575–579.

Conant, R. C., (1976). "Laws of Information Which Govern Systems". IEEE Trans. on SMC, SMC-6, 4, 240–255, April 1976.

Fahlman, S. E., Hinton, G. E., Sejnowski, T. J. (1983). "Massively Parallel Architectures for AI: NETL, THISLE and Boltzmann Machines". Proceedings of National Conference on AI, Menlo Park, CA.

Fu, K. S., (1971). "Learning Control Systems and Intelligent Control Systems: An Intersection of Artificial Intelligence and Automatic Control". *IEEE Trans on Automatic Control*, Vol. AC-16, No. 1, 70–72.

Geman, S. and Geman, D. (1984). "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images". IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. PAMI-6, pp. 721–471, November 1984.

Hayes-Roth, et al., (1983). *Building Expert Systems*. Addison-Wesley, New York.

Hinton, G. E., Sejnowski, T. J. (1986). "Learning and Relearning in Boltzmann Machines". pp. 282–317, in *Parallel Distributed Processing*, ed. D. E. Rumelhart and J. L. McClellan, MIT Press.

Hopfield, J.J. (1982). "Neural networks and physical systems with emergent collective computation abilities". Proc. National Acad. Sci., U.S.A: Vol. 79, pp. 2554–2558; April.

Jaynes, E. T., (1957). "Information Theory and Statistical Mechanics". *Physical Review*, 106, 4.

Kam, M., Guez, A. and Cheng, R. (1987). "On Stable Points and Cycles in Binary Neural Networks"; Proceedings of the IEEE Intl. Symp. on Intelligent Control, pp. 321–326, Philadelphia, PA, January.

Kirkpatrick, S., Gelatt, C.D. Jr., and Vecchi, M. P. (1983). "Optimization by Simulated Annealing"; *Science*, Vol. 220, Number 4598, May 13, 1983.

Kolmogorov, A. N., (1956). "On Some Asymptotic Characteristics of Completely Bounded Metric Systems". Dokl Akad Nank, SSSR, Vol. 108, No. 3, pp. 385–9.

Levis, A., (1988) "Human Organizations as Distributed Intelligence Systems". Proceedings 1st IFAC/IMACS Symposium on Distributed Intelligent Systems, Varna, Bulgaria.

Meystel, A., (1985). "Intelligent Motion Control in Anthropomorphic Machines", Chapter in *Applied Artificial Intelligence*, S. Andriole Ed. Pentrocellis Books, Princeton, NJ.

Saridis, G. N., (1977a). *Self-organizing Controls of Stochastic Systems*. Marcel Dekker, New York, New York.

Saridis, G. N. (1977b). "Expanding Subinterval Random Search for System Identification and Control". IEEE Trans. on Auto. Control, pp. 405–412, June.

Saridis, G. N., (1979). "Toward the Realization of Intelligent Controls". *IEEE Proceedings*, Vol. 67, No. 8.

Saridis, G. N., (1983). "Intelligent Robotic Control". *IEEE Trans. on AC-29*, 4.

Saridis, G. N., (1985a). "Control Performance as an Entropy". *Control Theory and Advanced Technology*, 1, 2.

Saridis, G. N., (1985b). "Foundations of Intelligent Controls". *Proceedings IEEE Workshop on Intelligent Controls*, p. 23, RPI, Troy, NY.

Saridis, G. N., (1988). "Entropy Formulation of Optimal and Adaptive Control", IEEE Transactions on AC, Vol. 33, No. 8, pp. 713–721.

Saridis, G. N. and Graham, J. H., (1984). "Linguistic Decision Schemata for Intelligent Robots", *Automatica*, Vol. 20, No. 1, 121–126.

Saridis, G. N., Stephanou, H. E., (1977). "A Hierarchical Approach to the Control of a Prosthetic Arm", *IEEE Trans. on SMC*, Vol. SMC-7, No. 6, pp. 407–420.

Saridis, G. N. and Valavanis, K. P., (1988). "Analytical Design of Intelligent Machines", *Automatica the IFAC Journal*.

Stephanou, H. E., (1986). "Knowledge Based Control Systems". *IEEE Workshop on Intelligent Control 1985*, p. 116, RPI, Troy, New York.

Vamos, T., (1986). "Metalanguages – Conceptual Models, Bridge Between Machine and Human Intelligence", Working paper E/37, Hungarian Academy of Science.

Winston, P. (1977). *Artificial Intelligence*. Addision-Wesley, New York.

Zames, G., (1979). "On the Metric Complexity of Causal Linear Systems, ϵ-entropy and ϵ-dimension for Continuous Time". *IEEE Trans. Automat. Control*, Vol. AC-124, pp. 222–230, April.

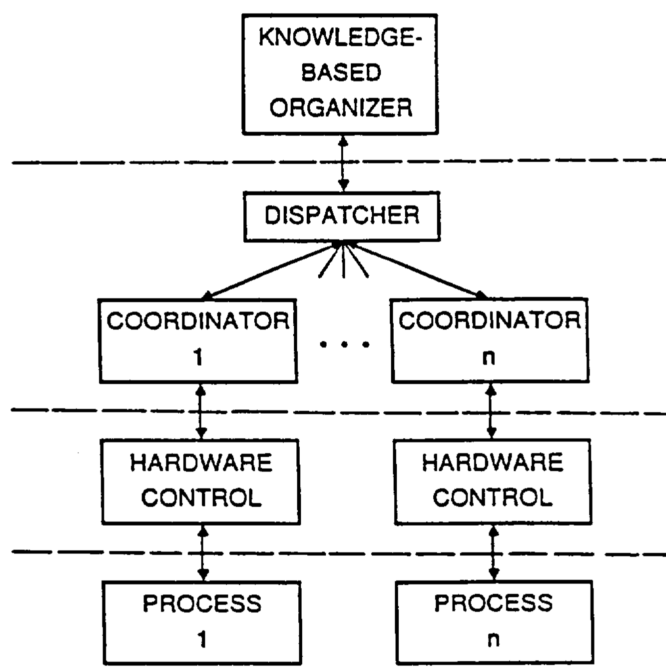

Fig. 1. Hierarchical Intelligent Control System

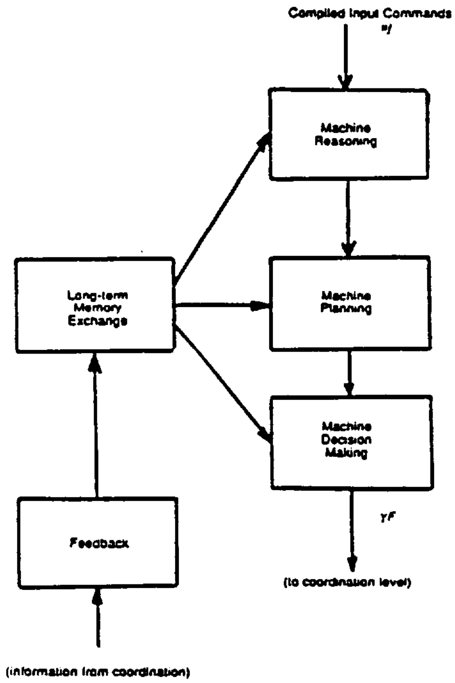

Fig. 2. Block Diagram of the Organization Level

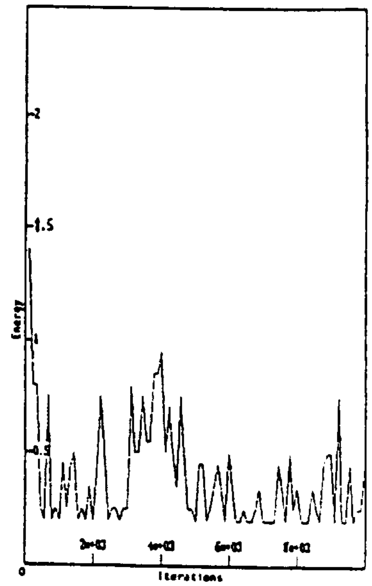

Fig. 3. Simulated Annealing: Nominal Case

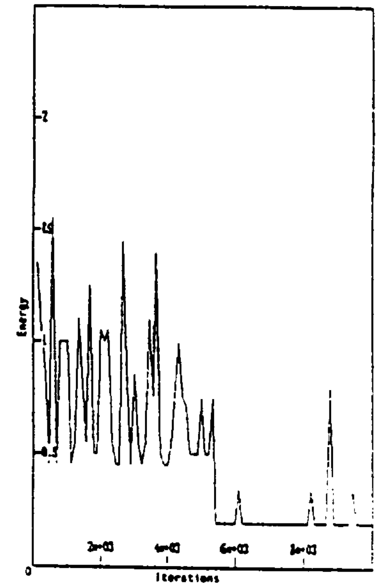

Fig. 4. Simulated Annealing: Second Case

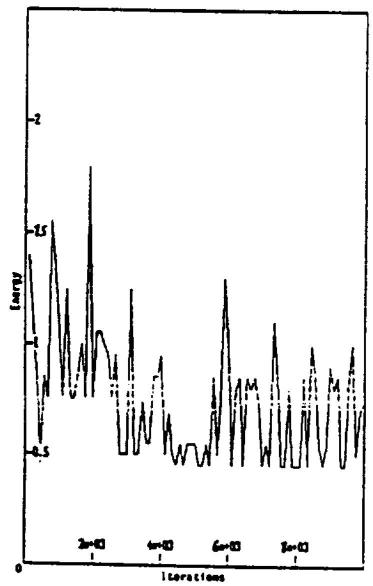

Fig. 5. Simulated Annealing Third Case

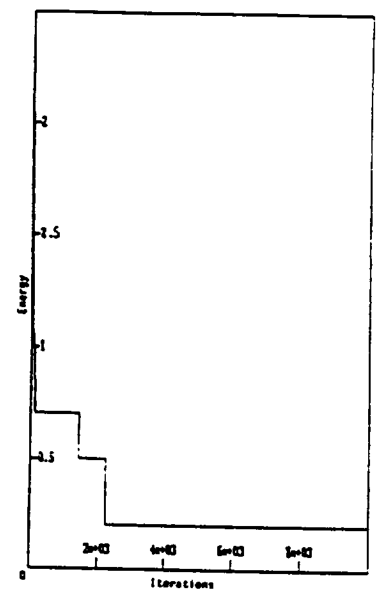

Fig. 6. ESRS: Nominal Case

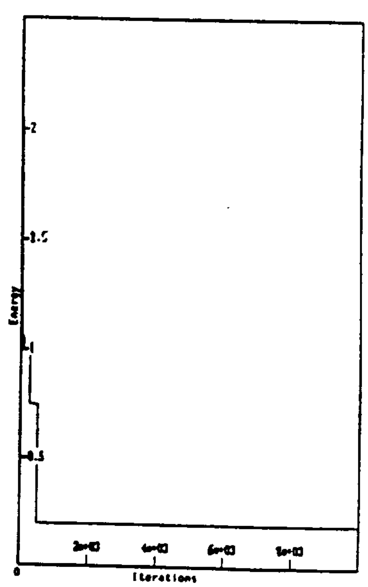

Fig. 7. ESRS: Second Case

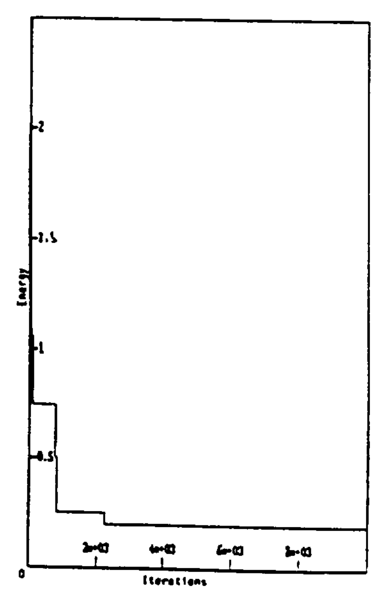

Fig. 8. ESRS: Third Case